Math 4527 (Number Theory 2)

Lecture #14 of 38 $\sim$ February 22, 2021

Elliptic Curve Factorization

- Properties of Order
- Elliptic Curve Factorization

This material represents §7.1.3-7.2.1 from the course notes.

The Group Law

For convenience in doing numerical computations, we can write down the general formula for the addition law on any curve:

Proposition (Explicit Group Law)

Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be points on the elliptic curve $E : y^2 = x^3 + Ax + B$. Then $P_1 + P_2 = (x_3, y_3)$ where $x_3 = m^2 - x_1 - x_2$ and $y_3 = -m(x_3 - x_1) - y_1$,

$$\text{with } m = \begin{cases} (y_2 - y_1)/(x_2 - x_1) & \text{if } P_1 \neq P_2 \\ (3x_1^2 + A)/(2y_1) & \text{if } P_1 = P_2 \end{cases}.$$

If m is infinite, then $P_1 + P_2 = \infty$.

Note that group law is rational, in the sense that the result is always a rational function of the inputs. In particular, the sum of two points whose coordinates lie in a field K will also lie in K.

Orders of Points on Elliptic Curves, I

Now that we've established some properties of the group law, we can use it to construct analogies between the structure of the points on an elliptic curve modulo p under addition and the units modulo n under multiplication.

- The point, so to speak, is that the points on an elliptic curve modulo p and the invertible residue classes modulo n are both finite abelian groups (E under the addition law, $(\mathbb{Z}/m\mathbb{Z})^{\times}$ under multiplication).

Orders of Points on Elliptic Curves, II

Our first goal is to define the order of a point on an elliptic curve. To do this we will use the addition operation on the curve:

Definition

Suppose E is an elliptic curve defined over a field K, and P is a point on E. For any positive integer k, we define the point kP to be the sum $\underbrace{P + P + \cdots + P}_{k \text{ terms}}$, and we also define $0P = \infty$ and $(-k)P$ as the additive inverse $-(kP)$.
The smallest positive k for which $kP = \infty$ is then called the <u>*order*</u> *of P; if no such k exists, then we say P has infinite order.*
A point of finite order is called a <u>*torsion point*</u> *and a point with $mP = \infty$ is called an* <u>*m-torsion point*</u>.

This is the same as the usual definition of the order of an element of a group, and the (m-)torsion elements of an abelian group.

Orders of Points on Elliptic Curves, III

A few remarks:

- Note that kP is well-defined because the addition law is associative: it does not matter the order in which we perform the additions. Likewise, we can see more or less immediately that $(a+b)P = aP + bP$ for any integers a and b.
- Over the real or complex numbers, "most" points on an elliptic curve will have infinite order.
- More precisely, as we will essentially show later, the set of torsion points on an elliptic curve over $\mathbb{C}$ is countably infinite, while the set of all points on the curve is uncountable.
- As we will show in a moment, however, on an elliptic curve modulo p all points have finite order.

Orders of Points on Elliptic Curves, IV

Example: Find the order of the point $P = (1, 3)$ on the elliptic curve $E : y^2 = x^3 + 4x + 4$ modulo 5.

Orders of Points on Elliptic Curves, IV

<u>Example</u>: Find the order of the point $P = (1, 3)$ on the elliptic curve $E : y^2 = x^3 + 4x + 4$ modulo 5.

- We simply compute the multiples of P using the addition law repeatedly.
- We obtain $2P = P + P = (2, 0)$, $3P = 2P + P = (1, 2)$, $4P = 3P + P = \infty$.
- Since $4P$ is the smallest multiple of P that gives the point ∞, the order of P is 4.

Orders of Points on Elliptic Curves, IV

We can compute large multiples of a particular point using successive doubling, in analogy to the procedure of successive squaring:

Algorithm (Successive Doubling Algorithm)

To compute kP, first find the binary expansion of $k = \underline{b_j b_{j-1} \cdots b_0}$. Then compute the multiples $2P$, $4P$, $8P$, ... , $2^j P$ by using the doubling part of the addition law. Finally, compute $kP = \sum_{\substack{0 \le i \le j \\ b_i = 1}} 2^{b_i} P$ using the addition law.

For example, to compute $77P$, we write $77 = 64 + 8 + 4 + 1$ compute $P, 2P, 4P, \ldots, 64P$ via doubling, and then add up $64P + 8P + 4P + P = 77P$.

Orders of Points on Elliptic Curves, V

The successive doubling algorithm is analogous to successive squaring inside $\mathbb{Z}/m\mathbb{Z}$.

- We can speed the successive doubling procedure up a bit by also using subtractions: unlike with modular arithmetic, where it is comparatively expensive to compute inverses, if $P = (x, y)$ then we have the trivial formula $-P = (x, -y)$.
- We will also observe that this procedure works for any elliptic curve, not just an elliptic curve modulo p. The only issue is that large multiples of a typical point will usually grow very complicated over an infinite field.

Orders of Points on Elliptic Curves, VI

Orders of points on an elliptic curve share many of the same properties as orders of units modulo an integer m, and the proofs of these results are also essentially the same.

Proposition (Properties of Order on Elliptic Curves)

Suppose E is an elliptic curve and P is a point on E.

1. *If P has finite order k and $mP = \infty$, then k divides m.*
2. *If $mP = \infty$ but $(m/q)P \neq \infty$ for any prime divisor q of m, then P has order m.*
3. *If E is an elliptic curve modulo a prime p and N is the number of points on E modulo p, then $NP = \infty$. In particular, the order of P divides N.*

Orders of Points on Elliptic Curves, VII

1. If P has finite order k and $mP = \infty$, then k divides m.

Proof:

- Suppose $mP = \infty$ and write $m = qk + r$ where $0 \le r < k$.
- We then have $rP = mP + (-qk)P = mP + (-q)(kP) = \infty + (-q)\infty = \infty + \infty = \infty$.
- Since $rP = \infty$ and $0 \le r < k$, the only possibility is to have $r = 0$: otherwise this would contradict the minimality of k. Thus $m = qk$ so k divides m.

Orders of Points on Elliptic Curves, VIII

2. If $mP = \infty$ but $(m/q)P \neq \infty$ for any prime divisor q of m, then P has order m.

Proof:

- Suppose the order of P is k. Then since $mP = \infty$, by (1) we know that k divides m.
- If $k < m$, then there must be some prime q in the prime factorization of m that appears to a strictly lower power in the factorization of k: then k divides m/q.
- But then $(m/q)P = \infty$ since m/q is a multiple of k, but this is contrary to the given information. Thus $m = k$ so P has order m.

Orders of Points on Elliptic Curves, IX

3. If E has a finite number N of points (in particular, if E is any elliptic curve modulo any prime p), then $NP = \infty$. In particular, the order of P divides N.

Remarks:

- This result is an analogue of Euler's theorem for $\mathbb{Z}/m\mathbb{Z}$.
- It is an immediate corollary of Lagrange's theorem from group theory (the order of any element of a group divides the number of elements in the group).
- In our case, we can give a self-contained proof by adapting the usual argument for proving Euler's theorem (which does, in fact, work for any finite abelian group).

Orders of Points on Elliptic Curves, X

3. If E has a finite number N of points (in particular, if E is any elliptic curve modulo any prime p), then $NP = \infty$. In particular, the order of P divides N.

Proof:

- Suppose the points on E are $Q_1, Q_2, \cdots, Q_N$ and consider the points $Q_1 + P, Q_2 + P, \cdots, Q_N + P$: we claim that they are simply the points $Q_1, Q_2, \cdots, Q_N$ again (possibly in a different order).
- Since there are N points listed and they all lie on the curve E, it is enough to verify that they are all distinct.
- So suppose $Q_i + P = Q_j + P$. Then we can write $Q_i = Q_i + \infty = Q_i + (P + (-P)) = (Q_i + P) + (-P) = (Q_j + P) + (-P) = Q_j + (P + (-P)) = Q_j + \infty = Q_j$, where we used associativity and the properties of ∞ and inverses. (Morally, we simply subtracted P from both sides.)

Orders of Points on Elliptic Curves, XI

3. If E has a finite number N of points (in particular, if E is any elliptic curve modulo any prime p), then $NP = \infty$. In particular, the order of P divides N.

Proof (continued):

- Thus the points $Q_1 + P, Q_2 + P, \cdots, Q_N + P$ are simply $Q_1, Q_2, \cdots, Q_N$ in some order.
- Adding up all the terms then yields $(Q_1 + P) + \cdots + (Q_N + P) = Q_1 + \cdots + Q_N$, and upon rearranging and subtracting $Q_1 + \cdots + Q_N$ from both sides (in the same way as above), we obtain $NP = \infty$ as desired.
- The second statement follows immediately from $NP = \infty$ and (1) above.

Orders of Points on Elliptic Curves, XII

Example: Show that the point $P = (1, 3)$ has order 15 on the elliptic curve $E : y^2 = x^3 + 4x + 4$ modulo 13.

Orders of Points on Elliptic Curves, XII

<u>Example</u>: Show that the point $P = (1, 3)$ has order 15 on the elliptic curve $E : y^2 = x^3 + 4x + 4$ modulo 13.

- It is a straightforward check that $15P = \infty$ using successive doubling: we compute $2P = (12, 8)$, $4P = (6, 6)$, $8P = (0, 11)$, $16P = (1, 3)$. Then $15P = 16P - P = (1, 3) - (1, 3) = \infty$.
- Furthermore, we can compute $3P = 2P + P = (3, 2)$ and $5P = 4P + P = (10, 2)$.
- Since neither of these quantities is ∞, we conclude that the order of P must be $\boxed{15}$.

Orders of Points on Elliptic Curves, XIII

If we can compute the orders of some points on E, we can often use that information in conjunction with the Hasse bound to determine the number of points on E without actually computing them all.

- In the example from the previous slide, we exhibited a point of order 15 on the elliptic curve $E : y^2 = x^3 + 4x + 4$ modulo 13. Thus, by our results on orders, the number of points on E must be a multiple of 15.
- By the Hasse bound, the number of points on E must satisfy $|N - 14| \leq 2\sqrt{13}$, yielding the inequality $6.78 \leq N \leq 21.22$. The only multiple of 15 in this range is 15 itself, so E must have exactly 15 points.

Orders of Points on Elliptic Curves, XIV

Example: Show that the point $P = (0, 2)$ has order 29 on the elliptic curve $E : y^2 = x^3 + x + 4$ modulo 23. Use the result to find the number of points on E and the group structure of E.

Orders of Points on Elliptic Curves, XIV

Example: Show that the point $P = (0, 2)$ has order 29 on the elliptic curve $E : y^2 = x^3 + x + 4$ modulo 23. Use the result to find the number of points on E and the group structure of E.

- It is a straightforward check that $29P = \infty$ using successive doubling and subtraction: we compute $2P = (13, 12)$, $4P = (1, 12)$, $8P = (14, 5)$, $16P = (8, 8)$, $32P = (11, 9)$. Then $3P = P + 2P = (11, 9)$ and so $29P = 32P - 3P = (11, 9) - (11, 9) = \infty$.
- Thus, the order of P is 29, as claimed.
- By the Hasse bound, the number of points on E must satisfy $|N - 24| \leq 2\sqrt{23}$, yielding the inequality $14.41 \leq N \leq 33.59$. The only multiple of 29 in this range is 29 itself, so E must have 29 points.
- Since 29 is prime, in fact the group structure is cyclic of order 29, and P (or any other nonidentity point) is a generator.

Elliptic Curve Factorization, I

Now that we have a reasonably good analogy between modular multiplication and the points on an elliptic curve modulo p under addition, we can use these analogies to develop algorithms for computational number theory and cryptography.

- We will first discuss how to use elliptic curve arithmetic to design an integer factorization algorithm (today).
- We then discuss how to develop several cryptographic protocols relying on the addition law on an elliptic curve. These will include a public-key cryptosystem based on ElGamal encryption, a key-exchange protocol based on Diffie-Hellman key exchange, and a digital signature algorithm.
- Since I'm not assuming you're intimately familiar with any of the $\mathbb{Z}/m\mathbb{Z}$ versions of these things, I will briefly review those as we go.

Elliptic Curve Factorization, II

We first explain how to create a factorization algorithm using elliptic curves based off of the method of Pollard's $(p-1)$-algorithm, as first proposed by Lenstra in 1985.

- In Pollard's $(p-1)$-algorithm, the basic idea is that if $n = pq$ and we choose a random integer a, then the order of a modulo p is likely to differ from the order of a modulo q.
- Thus, if the order of a mod p is k and the order of a mod q is bigger than k, then $a^k \equiv 1 \pmod{p}$ but $a^k \not\equiv 1 \pmod{q}$.
- Then $\gcd(a^k - 1, n) = p$. Thus, we can find a factorization of n by computing $a^k - 1 \bmod n$ (this is quick using successive squaring mod n) and then taking its gcd with n (also quick using the Euclidean algorithm).

Elliptic Curve Factorization, III

The nonobvious part is how to find an exponent k such that $a^k \equiv 1 \pmod{p}$ but $a^k \not\equiv 1 \pmod{q}$.

- We don't need to find the exact order of a mod p: any multiple of it will suffice, as long as that multiple is not also a multiple of the order of a mod q.
- A decent option that is also easy to implement is to evaluate the values $a^{1!}$, $a^{2!}$, $a^{3!}$, $a^{4!}$, ... , $a^{B!}$ modulo n (for some bound B), since the jth term is simply the jth power of the previous term.
- This procedure is guaranteed to return a result congruent to 1 modulo p provided that the order of a divides $B!$.

Elliptic Curve Factorization, IV

Algorithm (Pollard's $(p-1)$-Algorithm)

Let n be composite. Choose a bound B and a residue a modulo n. Set $x_1 = a$, and for $2 \le j \le B$, define $x_j = x_{j-1}^j \pmod{n}$. Compute $\gcd(x_B - 1, n)$: if the gcd is between 1 and n then we have found a divisor of n. If the gcd is 1 or n, start over with a new residue a.

- If $p|n$ and $p-1$ has only small prime factors, then the order of any element modulo p will divide $B!$ where B is comparatively small. On the other hand, if the other prime factors q_i of n are such that $q_i - 1$ has a large prime factor, it is unlikely that a randomly chosen residue will have small order modulo q.
- Thus, when we apply Pollard's $(p-1)$-algorithm to a composite integer $n = pq$ where $p-1$ has only small prime divisors, it is likely that the procedure will quickly find the factorization. (This is the reason for the algorithm's name.)

Elliptic Curve Factorization, V

Example: Use Pollard's $(p-1)$-algorithm with $a = 2$ to find a divisor of $n = 4913429$.

- We start with $a = 2$, so that $x_1 = 2$. We compute $\gcd(x_j - 1, n)$ for each value of j until we find a gcd > 1:

Value	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$
x_j	2	4	64	2036929	251970	3059995	1426887
gcd	1	1	1	1	1	1	2521

- After the 7th step, we obtain a nontrivial divisor 2521, giving the factorization $n = \boxed{2521 \cdot 1949}$.
- Observe that $2521 - 1 = 2520 = 2^3 3^2 5^1 7^1$ has only small divisors, and indeed 2520 divides 7! (so we were guaranteed to obtain it by the 7th iteration of the procedure).
- However, $1949 - 1 = 2^2 \cdot 481$ has a large prime divisor, so it would usually take $B = 481$ to find 1949 as a divisor.

Elliptic Curve Factorization, VI

The speed of Pollard's $(p-1)$-algorithm depends on the size of the largest prime divisor of $p-1$, which can vary quite substantially.

- If p is an odd prime, $p-1$ is clearly even, so the worst-case scenario is to have $n = pq$ where $p = 2p_0 + 1$ and $q = 2q_0 + 1$ with p, q, p_0, q_0 all prime and where p and q are roughly equal. In such a case, we would require $B \approx p_0 \approx \frac{1}{2}\sqrt{n}$ in order to find the factorization (unless we are lucky with a).
- It is also a rather involved analytic number theory problem to estimate the "expected" running time for the algorithm. In general, if we use a bound $B = n^{\alpha/2}$, then we would expect to have a probability roughly $\alpha^{-\alpha}$ of finding a factorization. When $\alpha = 1/2$ this says we would have about a 25% chance of obtaining a factorization if we take $B = n^{1/4}$.

Elliptic Curve Factorization, VII

Let's now construct an analogous procedure using elliptic curves:

- Again, suppose $n = pq$ is a product of two primes, and suppose we choose a (nonsingular) elliptic curve $E : y^2 = x^3 + Ax + B$ over the integers along with a point P on the curve.
- The order of P on E_p, the reduction of E modulo p, is unlikely to be exactly equal to the order of P on E_q, the reduction of E modulo q.
- If the order of P on E_p is k and the order of P on E_q is larger than k, then $kP = \infty$ on E_p but $kP \neq \infty$ on E_q.

Elliptic Curve Factorization, VIII

Now the question arises: how can we detect this behavior?

- In Pollard's $(p-1)$-algorithm, we performed all our calculations modulo n, so let's try that here: namely, doing all of our computations on the curve E_n, the reduction of the curve E modulo n, using the addition law formulas defined over the integers modulo n.
- Assuming that this reduction is well-defined, the addition law will still obey all of the requirements we put on it (namely, it will be commutative, associative, have an identity ∞, and have inverses).

Elliptic Curve Factorization, IX

However, the addition law formulas require a division when computing the slope of the line, and if this slope requires dividing by a nonzero number that is not invertible mod n, then we will not be able to evaluate the result.

- Just to be clear, if we were dividing by zero itself, then we would simply obtain a slope of ∞.
- The problem is that there is no sensible way to interpret (e.g.,) a slope of $1/2$ modulo 6.
- This may seem like it's a problem, but actually, it's exactly what we want!

Elliptic Curve Factorization, X

Specifically, suppose we obtain an "illegal" denominator when we do one of these calculations.

- This means that the slope of the line is ∞ modulo one of the prime divisors of n, but not ∞ modulo the other.
- We can use this information to factor n by taking the gcd of the problematic denominator with n.
- Another way to interpret this idea is using the Chinese remainder theorem: a point (x, y) lies on E_n if and only if it lies on the curve $E_p : y^2 = x^3 + Ax + B$ modulo p and the curve $E_q : y^2 = x^3 + Ax + B$ modulo q.
- Thus, the points on E_n can equivalently be thought of as pairs of points (P, Q) of points on E_p and E_q. We are then seeking to detect when a multiple of a pair (P, Q) is ∞ in one coordinate but not in the other.

Elliptic Curve Factorization, XI

Example: Examine what happens when trying to add the point $P = (1, 3)$ to the point $Q = (15, 4)$ on the elliptic curve $E_{21} : y^2 = x^3 + 4x + 4$ modulo 21.

Elliptic Curve Factorization, XI

Example: Examine what happens when trying to add the point $P = (1, 3)$ to the point $Q = (15, 4)$ on the elliptic curve $E_{21} : y^2 = x^3 + 4x + 4$ modulo 21.

- To find $P + Q$ we first compute the slope of the line joining them: it is $\frac{4-3}{15-1} = \frac{1}{14}$.
- However, this quotient is not defined modulo 21, since 14 is not relatively prime to 21.
- In this case, we see that $\gcd(21, 14) = 7$ is a proper divisor of 21: we have used this "failed" point addition to get a factorization of n.

Elliptic Curve Factorization, XII

Example: Examine what happens when trying to double the point $P = (1, 3)$ on the elliptic curve $E_{21} : y^2 = x^3 + 4x + 4$ modulo 21.

Elliptic Curve Factorization, XII

Example: Examine what happens when trying to double the point $P = (1, 3)$ on the elliptic curve $E_{21} : y^2 = x^3 + 4x + 4$ modulo 21.

- To find $2P$ we first compute the slope of the slope of the tangent line, which is $\frac{3(1)^2 + 4}{2 \cdot 3} = \frac{7}{6}$ by implicit differentiation.
- Just like before, this ratio is not defined modulo 21 since 6 is not relatively prime to 21, and just like before, $\gcd(21, 6) = 3$ is a proper divisor of 21.
- Ultimately, what is happening in the example from the last slide is that $P + Q = \infty \pmod 7$ but $P + Q \neq \infty \pmod 3$. Here, we see $2P = \infty \pmod 3$ but $2P \neq \infty \pmod 7$.

Elliptic Curve Factorization, XIII

Now we just have to organize all of this into an algorithm. Again, we take guidance from Pollard's $(p-1)$-algorithm.

- In Pollard's $(p-1)$-algorithm, we compute $\gcd(a^{d!}-1, n)$ for $1 \leq d \leq M$ (for some choice of bound M) until we obtain a gcd that is larger than 1.
- The analogous calculation on an elliptic curve is to try computing $(d!)P$ on an elliptic curve $E_n : y^2 = x^3 + Ax + B$ modulo n for $1 \leq d \leq M$, and checking if we obtain a denominator that has a nontrivial gcd with n in the denominator: if so, we get a factorization of n.
- The only remaining question is how to choose an elliptic curve E along with a point P. An easy way to generate a pair (E, P) is to choose the coordinates of $P = (x_0, y_0)$ along with the value A first, and then set $B = y_0^2 - x_0^3 - Ax_0$.

Elliptic Curve Factorization, XIV

This is precisely Lenstra's algorithm:

Algorithm (Lenstra's Elliptic-Curve Factorization Algorithm)

Suppose n is composite.
Choose a bound M, a point $P = (x_0, y_0)$, and an integer A.
Let E_n be the elliptic curve $y^2 = x^3 + Ax + B$ modulo n with B chosen so that P lies on E.
Set $Q_1 = P$ and for $2 \leq j \leq M$, define $Q_j = jQ_{j-1}$ (on E_n).

If at any stage of the computation the point Q_j cannot be computed, due to a necessary division by a denominator d which is not 0 modulo n but which is not invertible modulo n, then $\gcd(d, n)$ is a proper divisor of n. If a divisor is not found and Q_M is not ∞, increase the value of M and continue the computation. Otherwise, if $Q_M = \infty$, repeat the procedure with a new choice of P and A.

Elliptic Curve Factorization, XV

We have already done all of the legwork to show that this algorithm will succeed, and we have done a few "toy" examples already.

- The main question is: how efficient is elliptic curve factorization, and how well does it work in practice?
- We will analyze these questions next time, and also do some less trivial examples.

Summary

We discussed some properties of orders of points on elliptic curves.

We discussed how to use elliptic curves to do integer factorization.

Next lecture: Examples and analysis of elliptic curve factorization, elliptic curve cryptography.